PCT/SE 03/01122
08-09-2003

REC'D 16 SEP 2003
WIPO PCT

PA 1055033

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 20, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/418,547***
FILING DATE: ***October 15, 2002***

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

N. WOODSON
Certifying Officer

10-16-02 A1PROV

10/15/02 JC960 U.S. PTO

Attorney's Docket No. 8197-4PR

PATENT

J1050 U.S. PTO 60/418547 10/15/02

COVER SHEET FOR FILING PROVISIONAL PATENT APPLICATION (37 CFR §1.51(c)(1))

Box PROVISIONAL PATENT APPLICATION
Commissioner for Patents
Washington, DC 20231

Date: October 15, 2002

This is a request for filing a PROVISIONAL PATENT APPLICATION under 37 C.F.R §1.53(c).

Docket No.	8197-4PR
Type a plus sign (+) inside this box ☐	+

INVENTOR(s)/APPLICANT(s)

Name: Martin Korling
Address: Skuggvägen 12
SE-191 45 Sollentuna
SWEDEN

Name: Jan Lemark
Address: Lindhagensgatan 63
SE-112 43 Stockholm
SWEDEN

Name: Tormod Hegdahl
Moserabben 14
N-0752 Oslo
NORWAY

Name: Magnus Harnesk
Sköldvägen 35 B
SE-186 40 Åkersberga
SWEDEN

TITLE OF THE INVENTION *(280 characters maximum)*

FLEXIBLE BEARER CHARGING

ENCLOSED APPLICATION PARTS *(check all that apply)*

☒ Specification (Number of Pages 25)

☐ Drawing(s) (Number of Sheets __)

In re: Martin Korling et al;
Filed: Concurrently Herewith
Page 2

☐ Claims (Number of Claims)
(A complete provisional application does not require claims 37 C.F.R. §1.51(c)).

☐ Application Data Sheet. See 37 CFR §1.76

☐ Other:

CORRESPONDENCE ADDRESS

20792
PATENT TRADEMARK OFFICE

METHOD OF PAYMENT

☐ Applicant claims small entity status. See 37 CFR §1.27.
☒ Check or money order is enclosed to cover the filing fee of $160.00.
☐ Payment by credit card. Form PTO-2038 is attached.
☐ The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account No. 50-0220.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
☒ No.
☐ Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,

Scott C. Hatfield
Registration No. 38,176

CERTIFICATE OF EXPRESS MAIL

Express Mail Label Number EV 193629111 US
Date of Deposit: October 15, 2002
I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Box PROVISIONAL PATENT APPLICATION, Commissioner for Patents, Washington, DC 20231.

Joyce Paoli
Joyce Paoli

Flexible Bearer Charging

The basic question is: How to perform per-packet real-time charging in a packet switched communication system? Each packet should be charged differentially dependent on which service flow the packets belong to.

Performing differentiated rating on the packet level raises a fundamental challenge: rating is a complex process involving many input parameters (tariff plan, time and volume thresholds, subscriber profile, etc), while packet forwarding should be executed with lowest possible latency. Typical rating engines may not be optimised for packet forwarding and, vice versa, typical packet forwarding engines may not be optimised to perform complex pricing calculations.

The charging solution has to have the "real-time" propertity, meaning that the charging solution must be possible to used for so-called prepaid subscribers. Specifically, when a prepaid subscriber's (user's) credit account is empty, service execution (in this case packet forwarding) must be immediate affected.

Herein we will call the state-of-the-art solution a "multiple token bucket" solution. Such a solution consists of a control system and a packet forwarding system (see Figure 1).

The control system consists of a rating engine and credit accounts. The packet forwarding system contains multiple token buckets for every logged-in user.

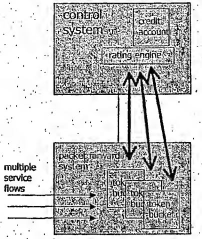


Figure 1 Multiple token bucket solution

When a user logs into the communication system, the packet forwarding system initates a control signalling sequence to the control system. The control system reserves an amount of credit towards the user's credit account. Let's call this amount the "credit reservation". The control system determines the set of services this user is allowed to use. The set of identifiers for these allowed services are sent to the packet forwarding system. For each service identifier, the packet forwarding engine initiates a resource reservation signalling sequence towards the control system. For each such sequence, the service is rated by the rating engine part of the control system, using a

tariff plan. The rating value is used to translate parts of the credit reservation into a "resource reservation" (typically data volume related), which is sent back to the packet forwarding system. Each resource reservation is put into a resource token bucket. Thus the packet forwarding system contains multiple token buckets, one for each allowed service.

When traffic is flowing through the communication system the packet forwarding system classifies each packet to determine which service it belongs to. Then the token bucket for that service is decremented.

When a token bucket is empty, the usage is confirmed towards the control system and a new resource reservation is done.

The multiple token bucket solution employs a separate resource reservation signalling sequence for each service in the set of allowed services. This may create a large amount of signalling traffic between the control system and the packet forwarding system. This may require high processing capabilities in the systems and high capacity transport between them.

Furthermore, the multiple token bucket solution may have drawbacks in the area of "over-reservation". This is illustrated by the following quantitative example: Assume that an end-user has a service mix according to the following list, including pricing information (note: cent could be euro-cents or dollar-cents, 10cents ~ 1SEK):

1. MMS, price 30cents per message, typically circa 50kB each, up to 1MB, typical "MMS session" contains 5 MMS messages

2. WAP surfing to "Internet", uncontrolled sites, price: 700cents per 1MB

3. WAP to controlled destination 1 (e.g., WSJ), price: 100cents per 1MB

4. WAP to controlled destination 2 (e.g., travel.com), price: 100cents per 1MB

5. WAP to controlled destination 3 (e.g., top-up and help pages), price: zero volume charge

In a multiple bucket system, this example gives rise to the following reservations at login (PDP Context Activation).

1. MMS: no credit reservation made, volume bucket contains 1MB for threshold purposes

2. WAP surfing to "Internet", uncontrolled sites: for example a bucket of 100kB translates into a credit reservation of 70cents

3. WAP to controlled destination 1(e.g., WSJ): for example a bucket of 50kB translates into a credit reservation of 5cents

4. WAP to controlled destination 2 (e.g., travel.com): for example a bucket of 50kB translates into a credit reservation of 5cents

5. WAP to controlled destination 3 (e.g., top-up and help pages), for example a bucket of 50kB, no credit reservation

Let's say that an end-user has 100cents left on the account. He logs in to send an MMS. At login, 80cents of credit is reserved. When he tries to send the MMS, the 30cent message charge is refused. This is an example of the "over-reservation" problem. If the operator wants to limit the total credit reservation toward the packet forwarding system (let's say 50 cents), this has to be translated into limits in the volume reservation defaults for each service. The problem gets worse as the number of services increases.

A proposed solution (see Figure 2) according to embodiments of the present invention can include a control system and a packet forwarding system.

The control system can include:

- a charging policy decision point (also called rating engine)
- a credit account function, which holds the users' credit accounts for real-time charging purposes

The packet forwarding system can include:

- a charging policy enforcement point
- a token bucket function per logged-in user

Examples of packet forwarding systems are: a switch, a router, a GGSN node in the GPRS/UMTS system, a PDSN node in the cdma2000 system, and a proxy.

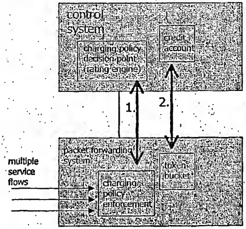


Figure 2 Single token bucket design with charging policy

When a user logs into the communication system, the packet forwarding system initates a control signalling sequence to the control system. The control system determines the set of services identifiers this user is allowed to use. The rating engine part of the control system calculates "charging policy" based on a tariff plan and other input data (for example, time-of-day, the user's roaming status, the aggregated transferred data volume for the user). This calculation is illustrated in Figure 3 and is called "dynamic pre-rating". A "charging policy" includes of a "user rating table" and set of validity conditions. The user rating table includes rating values for each "service class" the user is allowed to use. The "service class" concept is introduced to limit the size of the user rating table part of the charging policy. A "service class" is a family of services with the common property that they have exactly the same

charging pattern (tariff plan). The charging policy is sent back to the packet forwarding system.

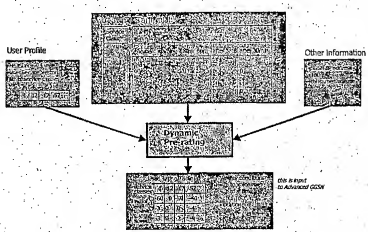
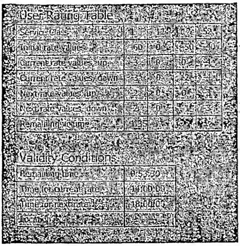


Figure 3 Process of calculating the "charging policy"

Next, the packet forwarding system initiates a reservation signalling sequence towards the credit account part of the control system. The credit account function reserves an amount of credit towards the user's credit account. Let's call this amount the "credit reservation". The credit reservation is sent back to the packet forwarding system and put into a user-specific token bucket.

When traffic is flowing through the communication system the packet forwarding system classifies each packet to determine which service it belongs to. Then the charging policy is used to determine which amount should be decremented from the token bucket.

The packet forwarding system continuously checks if the validity conditions are still satisfied. When that is no longer true, a signalling sequence is initiated to get a new up-dated charging policy. When a rate change should be performed at a certain time-of-day, T, the charging policy (user rating table part), contains the rate values that should be used both before and after T. Figure 4 shows an example of a charging policy.



User Rating Table

Service Class	1	[illegible]	[illegible]	52
Initial rate values	60	0	50	40
Current rate values, up	3	[illegible]	0	[illegible]
Current rate values, down	3	0	[illegible]	[illegible]
Next rate values, up	2	20	0	[illegible]
Next rate values, down	2	0	[illegible]	[illegible]
Remaining volume	135	[illegible]	[illegible]	[illegible]

Validity Conditions

Remaining time	0:53:30
Time for current rate	14:00:00
Time for next rate	18:00:00
Location	[illegible]

Figure 4 Charging policy, containing user rating table and validity conditions

When a token bucket is empty, the usage is confirmed towards the control system and a new resource reservation is done. In the confirmation signal, information about the usage of the different services is added explicitly.

According to embodiments of the present invention, a reduced amount of signalling traffic can be realized between the control system and the packet forwarding system.Reduced risk for over-reservation towards the credit account function can be realized by having a single token bucket that is shared by all service flows. Billing (post-paid) can be based on the real-time signalling, using the explicit service usage information carried by the confirmation messages. When a rate change should be performed at a certain time-of-day, T, the signalling load can be reduced by including, in the charging policy (user rating table part), both rates before and after T. When a user credit account becomes empty, it is possible to apply a fine-grained policy control of the traffic since the charging policy is there. The size of the charging policy can be reduced by introducing a "service class" concept.

According to embodiments of the present invention, a token bucket system can be provided for bearer charging in a packet switched communication system comprising a control system and a packet forwarding system. The packet forwarding system can include a token bucket function per logged-in user with a single token bucket per user. A charging policy decision point can reside in the control system and a charging policy enforcement point can reside in the packet forwarding system.

According to additional embodiments of the presen invention, a method to reduce control signalling for real-time per-packet charging in a packet-switch communication system can be provided. In addition, a method to reduce an over-reservation problem for real-time per-packet charging in a packet-switch communication system can be provided. A single token bucket per logged-on user (or per logical connection for a user) in the packet forwarding system can be provided. A charging policy can be provided expressing rating values for

different services (packet flows), rating values that are used to deduct tokens from the single token bucket, and validity conditions for these rating values.

Methods of applying such a charging policy to multiple simultaneous packet flows to determine the amount of tokens to be decremented from the single token bucket can also be provided. The concept of a "service class" can be provided as a method to reduce the size of a charging policy. The concept of letting the charging policy include rating values that are applicable both before and after a certain validity condition is met can be provided. In addition, methods can be provided to add information about service usage to the confirmation signalling towards to credit account function in the control system.

Multiple token bucket solutions to perform per-packet real-time charging in a packet switched communication system, where the packet should be charged differentially dependent on which service flow the packets belong to, may employ a separate resource reservation signalling sequence for each service in the set of allowed services. These solutions may require a large amount of signalling traffic and may have drawbacks in the area of "over-reservation".

Embodiments of the present invention may address one or more of these problems by introducing a control system including:

- a charging policy decision point (also called rating engine)
- a credit account function, which holds the users' credit accounts for real-time charging purposes

and a packet forwarding system including:

- a charging policy enforcement point
- a token bucket function per logged-in user

In the drawings and specification, there have been disclosed typical preferred embodiments of the invention and, although specific terms are employed, they are used in a generic and descriptive sense only and not for purposes of limitation. The following claims are provided to ensure that the present application meets all statutory requirements as a priority application in all jurisdictions and shall not be construed as setting forth the full scope of the present invention.

1. A token bucket system for flexible bearer charging comprising:
a control system including,
   - a charging policy decision point, and
   - a credit account function;

a packet forwarding system including,
   - a charging policy enforcement point, and
   - a token bucket function per logged-in user with a single token bucket per user; and

and a protocol for communication between a Control system and a Packet forwarding system.

60918547.101502

ERICSSON

Solution Description; Flexible Bearer Charging and Service Authorisation

Contents

1 Flexible bearer charging and control

GPRS operators may need to use service-based charging for operator-provided services and partner-provided services. These services should generally not be charged also by volume. Simultaneously, consumers will access services for which volume-related charging is the only possible (usage related) charging method.

Using several GPRS APNs for differentiating between different consumer service flows has disadvantages due to APN configuration management (terminal and network) as well as IP address handling in the terminal.

Therefore, there is a need for Flexible Bearer Charging, whereby service flows can be differentiated at the GPRS bearer level and volume charging can be applied in a flexible way. There is a need to support both prepaid and post-paid charging schemes.

There is a possibility to combine the Flexible Bearer Charging with operator-side Service Authorisation in order to control which services a subscriber can access.

2 Ericsson solution architecture

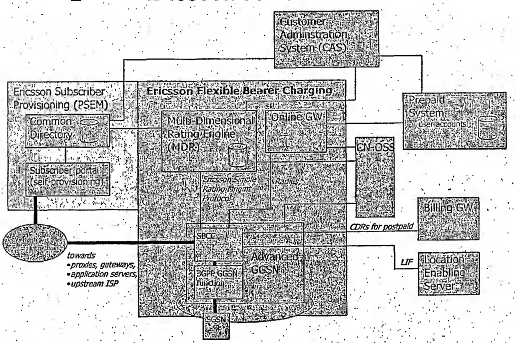


Figure 1 Solution architecture

The Ericsson Flexible Bearer Charging solution consists of the following components:

- The **Advanced GGSN**, which includes the Service-Based Charging and Control (SBCC) feature with a prepaid reservation engine.
- The **Multi-Dimensional Rating Engine**, which lets the operator define services and tariff plans, and dynamically provides the Advanced GGSN with rating information (cf. separate product description [MDR]).
- The **Online Gateway**, which provides an integration point for interworking towards the operator's prepaid system by means of flexible interface options. In the case when the pre-paid system supports a Radius interface, the Advanced GGSN can interfaces directly to it.
- The O&M solution integrated with the **Core Network OSS**, CN OSS.

If the operator wants to employ service authorisation (sometimes called service selection), Ericsson can provide the Subscriber Provisioning system PSEM (cf. Section 9), including the components:

- Common Directory server, for storage of user profile and service definition data
- Subscriber portal, for allowing the subscribers to self-manage their service profiles

The solution interfaces the following operator system components:

- The Customer Administration System system, where primary subscriber information is managed and stored
- The Prepaid system, where the subscribers' credit accounts are kept
- The Billing-GW, which is the gateway to the billing system
- The Location Enabling Server (LES)

3 Basic terminology

The following terminology is important knowledge to understand the Flexible Bearer Charging solution:

- Service, is anything that the Operator offers to its Subscriber that involves packet traffic.
- Service Identifier, is the unique identifier of a Service.
- Service Class, is the unique identifier of an arbitrary group of Services that the Operator wants to treat in the same way regarding charging. A Service can only belong to one Service Class.
- Service Filter, is specific filter setting that will handle a specific Service the Service Identifier is used to map the Service to its Service Filter.

This terminology is further described in the chapter Data structure and definitions

4 Main features

4.1 Differentiated Packet Rating

Performing differentiated rating on the packet level raises a fundamental challenge: rating is a complex process involving many input parameters (tariff plan, time and volume thresholds, subscriber profile, etc), while packet forwarding should be executed with lowest possible latency. Typical rating engines may not be optimised for packet forwarding and, vice versa, typical packet forwarding engines may not be optimised to perform complex pricing calculations.

To get the best of both worlds, Ericsson's solution employs a **two-stage rating process.**

1. **Dynamic Pre-rating,** which is performed by an external rating engine such as the Ericsson Multi-Dimensional Rating (MDR) product.
2. **Real-time packet rating** performed by the SBCC function in the Advanced GGSN.

The Dynamic Pre-rating process calculates a set of "temporary" rating values for each Service Class that a subscriber can use. This calculation makes full use of the tariff plan and the current subscriber situation (roaming status, time-of-day, etc). The rating values have a limited lifetime and have to be renewed for example when a tariff threshold is reached.

The Real-time packet rating is then performed by classifying a packet as belonging to a certain Service Class and using the "temporary" rating values for that packet. This can be done with very small delay in the forwarding process.

For example, we have the important case of employing zero volume-charge for services that have service-based charging (e.g., MMS). For such services, the Dynamic Pre-rating process would result in a "temporary" rating value of zero. The Real-time packet rating will then use the zero value and thus incur no prepaid charge for that traffic.

4.2 Prepaid Token reservation

The Advanced GGSN's SBCC function includes a reservation engine that interacts with the operator's prepaid system. Tokens are reserved and confirmed when consumed. A Token bucket is maintained per user and changed according to the rating value defined for the Service Class by the Real-time packet rating.

The Advanced GGSN uses the Radius protocol for reservation and confirmation. To get flexible interworking, supporting a variety of optional interfaces, the Ericsson product, Online-GW, is used as a gateway between Radius and for example Parlay.

4.3 Service-Based Charging and Control (SBCC)

The Advanced GGSN's SBCC function classifies packets using Service filtering based on IP header information and higher-layer protocols. While the solution architecture supports general stateful inspection and identification of URIs, initial focus is put on detecting MMS traffic over WAP 1.x and 2.x. After a packet has been classified as belong to a certain Service and Service Class, the Real-Time packet rating determines the rating value to change the Token bucket by.

The CDR generation is enhanced with per Service counter identified by the Service Identifier, so that CDRs can be used differentially towards Subscriber billing or settlement towards third-party service provider.

The Advanced GGSN also enforces certain policies when the Subscriber's prepaid account runs empty. Depending on roaming conditions, PDP Contexts can be de-activated or non-zero-rated packet can be stopped, while zero-rated traffic is let through (for example allowing access to top-up pages).

5 Major usage cases

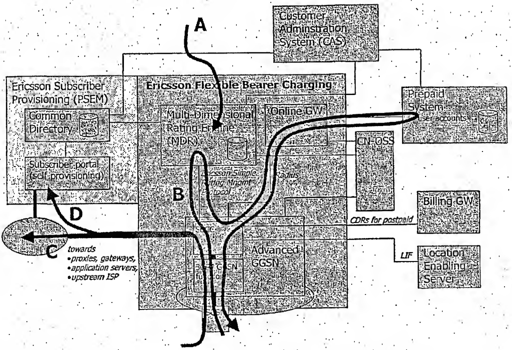


Figure 2 Major usage cases

Figure 2 illustrates how the Ericsson Flexible Bearer Charging solution works in four major usage cases: A, B, C, and D.

A **Provisioning of new services, products, and tariff plans.**
The Ericsson Multi-Dimensional Rating Engine provides the operator with a rich Graphical User Interface for managing products and tariff plans (cf. separate product description [MDR]).

B **User login (PDP Context Activation).**
When a GPRS subscriber connects (activates a PDP Context), the Advanced GGSN provides the Multi-Dimensional Rating Engine with the subscriber identity (and additional information) and requests rating information. The Multi-Dimensional Rating Engine performs the Dynamic Pre-rating calculation using the tariff plan, user profile, and other data. The Multi-Dimensional Rating Engine answers back with the rating information. Then the Advanced GGSN performs a Token reservation towards the prepaid system, via the Online-GW. The received Token quota is put in a Token bucket.

C **Forwarding of user traffic.**
The Advanced GGSN classifies packets into Service Classes and Service Identifiers using the Service Filters. The size of the packet is counted separately per Service for CDR generation. The Service Identifier is used in the CDR to identify the Service. A

rating value is determined based on the Service Class and the corresponding volume price (which could be negative, zero or positive) is deducted from the Token bucket. When the Token bucket is empty, a new Token request is send to the prepaid system. [JS Doesn't this need to be done when a threshold has been reached rather than when the bucket is empty, to avoid the necessity of passing packets which are un-rated?] If the prepaid account is empty, the Advanced GGSN can be configured to enforce a policy: either de-activation of PDP Context or just blocking non-zero-rated packets.

D **Subscriber self-provisioning.**
This is optional if the operator desires to employ service authorisation.

6 Data structure and definitions

6.1 Service and Service Class

Service could be anything that the Operator offers the Subscriber that involves packet traffic. A Service is defined by a unique Service Identifier. To perform the traffic-related parts of the FBC solution, which are handled by the SBCC, a Service Filter (SF) is used for each Service, see Figure 3. The mapping between the Service Filter and the Service is done by the Service identifier. In each Service Filter, a destination or source (depending on the traffic direction) IP addresses (and mask), TCP and UDP port numbers (ranges), and protocol numbers can be set. In order to handle inspection of higher protocol layers, a Service Filter can include a pointer to a Protocol Inspection Filter (PIF), which perform stateful inspection on packets.

The PIFs have specific configuration data (see Figure 4), for example containing URL/URIs. For packets that fit to the defined PIF rules Dynamic Filters are established to perform the actual filtering.

Finally, Service Filters and PIF configuration lines contain the Service Class number that represent the result of a filter matching an incoming packet. The Service Class number determines what rating will be applied to the packet.

Service Filter and PIF configurations are done in the Advanced GGSN on a per APN basis.

The benefit for the Operator of this two-stage approach is that if throughput and latency are most important, Services can be defined using only the Service Filter not the PIF. But for other Services the PIF functionality caters for a variety of possibilities that can be expanded later as new requirements appears. This approach avoids the overhead of stateful packet inspection where it is not required.

Service Filters

Service Identifie	filter rule								
	filter prio	IP address/mask	TCP/ UDP	port#	proto col#	ICMP type	protocol	PIF pointer	Service Class
1	1	100.18.0.0/16	*				WSP	1	
2	2	100.18.0.0/16	*				HTTP	2	
3	3	100.12.0.0	*						22
	4	*	*						60

Figure 3 Service Filter table, configured into the Advanced GGSN

Protocol Inspection Filter (PIF) configuration

PIF	Identifier	Service Class
1	mms1.oper.net	14
1	mms2.oper.net	14
1	*	15
2	music.com	52
2	*	53

Figure 4 PIF configuration

Example

Figure 3 and Figure 4 show an example where the operator's WAP GWs are placed at the subnet 100.18.0.0/16. Let's assume that these WAP GWs can also function as HTTP proxies. The Service Filters 1 and 2 both match packets to and from that subnet. Filter 1 is set to match the transport protocol WSP and invoke PIF number 1, while filter 2 will match the transport protocol HTTP and invoke PIF number 2. Assume that an incoming packet is a WAP-packet leading to a match for WSP. PIF 1 now employs the identifier data in the PIF configuration, in this case the URL/URI. In this example, it checks for the domain names of the MMS-centers of the operator. If the packet's WSP-header contains any of these domain names the resulting Service Class is 14. If the packet contains any other URL (the wild card *), the packet is non-MMS WAP traffic and is classified as Service Class 15.

Referring back to Figure 3, Service Filter 3 is set to match the subnet 100.12.0.0/16, which for example could be a partner service provider. All packets to and from that service provider will be classified as Service Class 22.

Finally, packets matching the wildcard rule (service filter 4), representing "other traffic", will be classified as Service Class 60.

6.2 User Profile and User Service Class Vector

A User Profile (cf. Figure 5) includes the subscriber identifier together with user specific data. The user data that is specific for the Flexible Bearer Charging Solution is the User Service Class Vector. The User Service Class Vector is a list of Service Classes (numbers) that the user is allowed to use. Note: if the Operator does not want to employ service authorisation all Service Classes should be listed in the User Service Class Vector.

User Profile

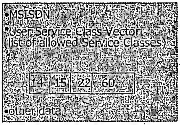


Figure 5 User profile

Continuing with the concrete example described in the previous section, the User Profile in Figure 5, allows the user to access Service Classes 14 (MMS traffic), 15 (other WAP traffic), 22 (the partner service provider), and 60 other traffic (general Internet access).

The User Profiles are managed by the Customer Relation Management system and could be stored in the Common Directory. Independent of where the User Profiles are stored they are replicated down to the Multi-Dimensional Rating Engine.

6.3 Tariff Plan, User Rating Table, and Validity Conditions

The Dynamic Pre-rating is done according to the Tariff Plan. The Tariff Plan is a method that, given parameters along several rating dimensions, produces a series of rating values.

The rating dimensions can be:

- roaming status (home or away)
- Time-of-Day (including month, year, etc)
- aggregated volume, (also called "transfer history")
- aggregated connect time
- QoS of the PDP Context
- geographical position
- special events

Figure 6, shows a simplified Tariff Plan where (referring back to the concrete example in previous sections) Service Class 14 (MMS traffic) is zero-rated except when the Subscriber is roaming, Service Class 15 (other WAP-traffic) is rated –2 units per byte. Traffic to and from the partner service provider (Service Class 22) is free-of-charge except when roaming.

These rating dimensions are independent and can be combined, so that a specific rate would apply, for example, between 6pm and 6am, above 3 MBytes, when roaming. The Tariff Plan in Figure 6 is just a schematic simplification of the rich, graphical capabilities of the Ericsson Multi-dimensional Rating Engine.

Tariff plan

Service Class	Tariffs				
	home	roaming	after 6pm	after 3MB	after 30min
14	0	-1	0	0	0
15	-2	-2	-2	-2	-2
22	0	-1	0	0	0
52	-1	-1	-1	-1	-1
60	-3	-4	-2	-2	-2

Figure 6 Simplified Tariff Plan example

For each Service Class in the User Service Class Vector, the Dynamic Pre-rating calculation generates five rating values: an "initial rate value", two "current rate values" and two "next rate values". For further information on the interaction between the Tariff Plan and the User Service Class Vector see section 7.2 Multi-Dimensional Rating. The "initial rate value" is

equal to the Token that should be deducted when either a Service Class is used for the first time or for the first used Service Class depending on a configuration value for that Subscriber. Thus initial fee can be based on Services or Subscriber. To handle different setting of what is meant by "first time" e. g. per day or month the usage of the "initial rate values" are send back to the Multi-Dimensional Rating Engine to be used for the next Dynamic Pre-rating calculation. The "current rate values" is the Token per byte to be charged for packets belonging to a Service Class that shall be used immediately see validity conditions below. The reason to have two is to allow different rating values for up and down link traffic, respectively. The "next rate values" shall be used after the expiration of the "current rate values" (cf. about validity conditions below). The purpose of the "next rate values" is to manage mass update due to tariff change, e.g., when going from peak hours to low traffic in the evening.

All these rating values are listed per Service Class to form the User Rating Table (see Figure 7).

User Rating Table

Service Class	1	12	17	52
Initial rate values	-60	0	50	-40
Current rate values, up	-3	0	0	-4
Current rate values, down	-3	0	2	-4
Next rate values, up	2	0	0	3
Next rate values, down	-2	0	-1	-3

Validity Conditions

Remaining volume	135
Remaining time	0:53:30
Time for current rate	14:00:00
Time for next rate	18:00:00
Location	

Figure 7 User Rating Table

The User Rating Table has a limited lifetime, specified by a family of **validity conditions**, which are calculated using the tariff thresholds for the dimensions:

- aggregated volume (also called "transfer history")
- aggregated connect time
- ToD
- geographical location

It is the responsibility of the SBCC to request an updated User Rating Vector when the threshold volume/time/location is reached.

7 Component functional description

7.1 SBCC in Advanced GGSN

Figure 8 shows an overview block diagram of the Service-based Charging and Control (SBCC) function, which resides in the Advanced GGSN.

The Service Filter and PIF configurations are configured in the SBCC using the CN-OSS.

The SBCC performs packet inspection according to the Service Filters. It invokes Protocol Inspection Filters (PIFs) for analysis of higher-layer protocols.

The SBCC has an extended set of traffic counters, so that CDRs can be generated differentially towards the billing system (for post-paid consumers) or for settlement towards third-party service provider.

For pre-paid subscribers, the SBCC utilizes the dynamic pre-rating and the reservation mechanisms. The SBCC distinguishes between post-paid and pre-paid subscribers either by using the charging characteristics information or by analysis of the subscriber IMSI. The SBCC can be configured to retrieve the User Service Class Vector also for post-paid subscribers. This should be done if the operator wishes to perform operator-side service authorization (cf. Section 9).

The SBCC function requests User Rating Tables from the Multi-Dimensional Rating Engine. This is triggered either by PDP Context Activation or that the validity conditions (for the User Rating Table) are no longer satisfied. When a new User Rating Table is requested, the SBCC reports back the transferred volume and the connect time as well as the status of the "initial rate values".

The SBCC function makes reservations towards the prepaid system, via the Online-GW, and thereby fills up each Subscriber's local Token bucket (kept in the SBCC). The amount of reservation and valid time are configurable in the SBCC. It down-counts each subscriber's Token bucket according to the User Rating Table.

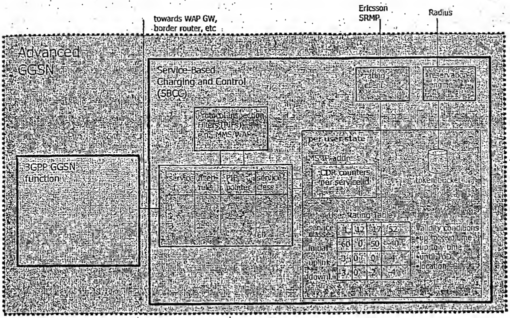


Figure 8 SBCC functional block diagram

A "one-time initial charge" is supported in two ways: First by Service Class, for each Service Class in the Subscriber's User Service Class Vector, the arrival of the first packet classified to that Service Class implies a deduction of the initial rate value from the Token bucket. The SBCC then puts zero into the initial rate value for that specific Service Class. Second by Subscriber, all Service Classes in the Subscriber's User Service Class Vector includes the same value. When a packet arrives for any of those Service Classes the Token bucket is decreased with the value and all values are set to zero. The initial rate value is reported as used back to the Multi-Dimensional Rating Engine, which then can give a zero initial rate value at subsequent requests.

For location-dependent rating the Advanced GGSN supports both forwarding of the SGSN IP address and a LIF interface to a LES. The Advanced GGSN can thus check the validity condition concerning geographical location and request a new User Rating Table in case the condition is no longer satisfied. For time dependent rating the SBCC supervises the validity time condition and requests a new User Rating Table when the threshold is passed. In addition, the SBCC supports maximum time subscriptions, i.e., connect up to 24 hours with one fee. This is managed by the remaining time validity condition, the SBCC counts down this value and if reaching zero, a new User Rating Table is requested. If a PDP context de-activation occurs before the value has become zero the value is send back to the Multi-Dimensional Rating Engine which could then update the aggregated time for that Subscriber.

For volume dependent rating the SBCC supervises the validity volume condition and requests a new User Rating Table when the threshold is passed. The SBCC decreases this value with the amount of traffic that passes and if reaching zero a new User Rating Table is requested. If a PDP context de-activation occurs before the value has become zero the value is send back to the Multi-Dimensional Rating Engine which could then update the aggregated volume for that Subscriber.

Service Control 1: What happens when the prepaid account is empty?

The SBCC architecture supports two "policy modes": "hard" and "home-liberal".

"Hard" mode

- PDP Context Activation denied if account empty
- Mid-session: if account empty then de-activate PDP Context

"Home-liberal" mode

- PDP Context Activation denied if account empty and roaming status = away
- PDP Context Activation allowed if account empty and roaming status = home
- Mid-session: if account empty and roaming status = away then de-activate PDP Context
- Mid-session: if account empty and roaming status = home, still allow packets for services classes with zero or "bonus" rating, but discard packets for services classes with non-zero rating.

In addition to these policies, the architecture allows for URI re-direction and a more generic event generation. URI re-direction is meaningful in situations when there is a browser running in the terminal and can then be used to let the Subscriber re-fill the prepaid account (top-up page). But in situations when there is no browser, it's useful to initiate a generic event, so that an SMS or MMS can be sent to the user.

The Advanced GGSN gets information about which policy to use along with roaming status from the Multi-Dimensional Rating Engine.

Service Control 2: What happens with a packet that has been classified as belonging to a Service Class that is not included in the subscriber's User Service Class Vector?

Packets will be silently discarded.

Service Control 3: How are packets treated which are received while waiting for a quota reservation reply from the prepaid system?

Packet will be forwarded anyway.

7.2 Multi-Dimensional Rating Engine

Figure 9 visualises how the Multi-Dimensional Rating Engine produces a User Rating Table by combining the User Profile (User Service Class Vector), the Tariff Plan, and additional information (such as roaming status, aggregated volume and connect time (stored in Multi-Dimensional Rating Engine database), time of day, and geographical location

In addition to the User Rating Table, a corresponding set of validity conditions are generated using tariff thresholds.

The User Rating Table and the validity conditions are provided to the Advanced GGSN.

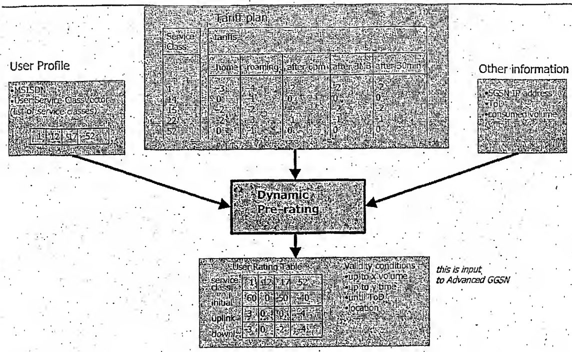


Figure 9 Dynamic Pre-rating, performed by the Multi-Dimensional Rating Engine

More specifically, the Multi-Dimensional Rating Engine performs the following procedure:

- accept request for a User Rating Table; input is the user identifier (MSISDN), SGSN IP-address, QoS of the PDP Context, etc.
- get the user's user profile (from a database, possibly the Directory)
- for each Service Class in that Subscriber's User Service Class Vector, get the service definition
- get the aggregated volume (transfer history) and aggregated connect time for this user (internal database in Multi-Dimensional Rating Engine)
- calculate roaming status based on SGSN IP address
- for each Service Class, calculate the relevant rating values using the tariff plan and information about the user's roaming status, aggregated volume and connect time, ToD, and geographical location
- calculate the validity condition for these rating values using the tariff thresholds in the dimensions: roaming status, aggregated volume and connect time (stored in Multi-Dimensional Rating Engine database), ToD, and geographical location
- construct the User Rating Table
- send the User Rating Table and the validity conditions to the Advanced GGSN

The high-availability features of the underlying platform provide redundancy for the Multi-Dimensional Rating Engine. In addition, multiple physical boxes can be deployed.

7.3 Online-GW

The Online-GW essentially performs protocol conversion between Radius (on the GGSN side) and another protocol towards the prepaid system. The architecture allows for TCP/IP based protocols such as XML, and SOAP.[JS and Corba?]

An SS7 interface can be discussed on customer request.

Redundancy as well as load balancing of the Online-GW is provided by multiple physical boxes.

7.4 Centralised O&M

The O&M for the Flexible Bearer Charging solution will be provided by the existing Core Network OSS. CN-OSS already provides extensive management support for the GGSN nodes this includes:

- Configuration management (network interfaces, application data, APN data)
- Fault management (collection mediation and presentation of all node alarms)
- Performance management (collection, storage and reports).

For the Online Gateway and Multi-Dimensional Rating Engine CN-OSS will provide basic support this includes:

- Fault Management (collecting, storing and presenting alarms from the nodes in one tool)
- The nodes will be included in the core network model (can be seen in the topology).
- CN-OSS will provide access to the nodes (remote) where the user can make any required changes using a CLI.

All Configuration of the MDR and Online-GW can be performed by using the remote access from CN-OSS.

It is not seen that any Performance data support is required for these two nodes. However CN-OSS provides an infrastructure for creation of reports related to this specific solution and also incorporation of other related performance data.

8 Interfaces

8.1 GGSN-Multi-Dimensional Rating Engine

This interface employs the Ericsson Simple Rating Management Protocol (SRMP), which is socket-based client-server protocol, carried on TCP/IP.

8.2 GGSN-Online-GW

The Radius protocol is used with specific extensions to support Token reservation and confirmation.

8.3 Online-GW – prepaid system

See Section 7.3.

8.4 O&M interfaces

Connection to CN OSS is provided by IP based protocols such as FTP, SNMP, HTTP.

9 Service Authorisation and Subscriber provisioning (PSEM)

As on option, Ericsson can supply the Personal Service Environment Manager (PSEM), which is the single point of Provisioning from both the Customer Administration System (CAS) as well as the End user. The end user can do self-provisioning from a web application, e.g., on a portal. PSEM also contains the Common Directory (CD). For service provisioning a state of the art SOAP/XML (CAI3G) interface is provided towards the CAS system.

PSEM contains the Ericsson Multi Activation Product (EMA) containing programmable interfaces towards various types of network elements e.g. a prepaid system.

The CD is the central repository, which contains common data such as the attributes of end-users and services and references to the location where the enabler/application specific information is stored.

The Common Directory contains a unique Ericsson Developed Data model. This model is non-static, meaning that it can be used towards affiliated data models already existing in the operators network. This inbuilt flexibility ensures fast integration in a mixed environment.

Applications can either read Directory Data directory via LDAP or use an API bundled with PSEM to retrieve the same information without the knowledge of the data model.

Specifically, if the operator wishes to perform operator-side service authorisation, the User Service Class Vectors are used to control which service classes a subscriber can access. The User Service Class Vectors are provisioned through the PSEM system. This allows for self-provisioning via the subscriber portal, which can also form the basis for top-up pages.

10 Benefits of the Ericsson Flexible Bearer Charging solution

By integrating the Flexible Bearer Charging in the Advanced GGSN, the operator may gain the following benefits

- reduce cost by reducing the number of network elements needed
- higher potential to utilize the GPRS-specific information and functions available in the Advanced GGSN, for example, PDP Context control, subscriber data, and QoS characteristics
- avoiding introducing interfaces between different network elements, and thereby reducing the integration cost

The Ericsson's Flexible Bearer Charging solution may include one or more of the following characteristics:

- The introduction of Service Class concept including the User Service Class Vector per Subscriber.
- The introduction of Dynamic Pre-rating at Subscriber connection including the User Rating Table and validity condition concepts.
- The introduction of Real-time packet rating based on Service Class.
- The Token bucket handling per Subscriber instead of per Service and the reservation at Subscriber connection.
- The introduction of Service Filter and Protocol Inspection Filter concepts.

10.1 Service Class concept

The Service Class concept allows the Operator to group related Services into a number of groups. The relation between the Services could be arbitrary except that the tariff must be the same since the rating is based on the Service Class. Still, each Subscriber could have different tariffs for a particular Service Class, but for a particular Subscriber, all Services in a Service Class will have the same rating (at a particular time). The benefit besides the Service grouping is that the amount of data per Subscriber is dramatically reduced since the number of Services could be in the order of thousands but the number of Service Class is up to 64. The Service Class concept is also the basis for operator-side service authorisation, which is an optional feature of the Flexible Bearer Charging solution.

10.2 Dynamic Pre-rating at Subscriber connection

The Dynamic Pre-rating at Subscriber connection means that all applicable rating values including their validity conditions are calculated before the Subscriber starts to use the different Services. This means that the delay when starting to use a Service (for example when retrieving MMS messages) can be reduced to a minimum leading to better end-user satisfaction and higher data throughput. It also reduces the control signalling traffic thus reducing the load of the Multi-Dimensional Rating engine. Both traffic data throughput enhancement and control signalling traffic reduction can lead to fewer physical boxes in the network. A high level of rating flexibility can be maintained as the rating values have well-specified validity and can be re-newed.

10.3 Real-time packet rating based on Service Class

The Real-time packet rating means that all packets belonging to a certain Service Class are rated in accordance with the rating value associated with that Service Class in real-time without need for control signalling towards the Multi-Dimensional Rating Engine. This means that the traffic delay while using a Service can be reduced leading to higher traffic data throughput. Both traffic data throughput enhancement and control signalling traffic avoidance can lead to fewer physical boxes in the network. Also, with this concept, we can avoid keeping track of service-level events. Such events should be charged at the application level, e.g., by the MMS server. Finally, we can avoid a full rating value calculation for each packet.

10.4 Token bucket handling per Subscriber

The Token bucket handling per Subscriber concept can reduce the distribution of Token reservations in the network thus reducing the risk to empty the prepaid account due to many different reservations. It also reduces the control signalling between the Advanced GGSN

and the Online Gateway and eventually the Pre-paid System thus reducing the needs of physical boxes for all three functions.

10.5 Service Filter and Protocol Inspection Filter concepts.

The Service Filter and Protocol Inspection Filter concepts allow the Operator to specify filter rules for a Service that can increase the traffic data throughput by just assigning rules in the Service Filter, while still having the possibilities to invoke more advance filter rules when necessary to distinguish the Service. The concepts as such are general which allow for further inclusion of new rules when needed for other Services.

Attorney Docket No. 8197-4PR

GLOSSARY OF TERMS

GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
PDSN	Packet Data Serving Node
CDMA	Code Division Multiple Access
GPRS APN	GPRS Access Point Node
MDR	Message Detail Recording
CN OSS	Core Network Operations Support System
MMS	Multimedia Messaging Service
IP	Internet Protocol
URI	Universal Resource Identifier
WAP	Wireless Application Protocol
CDR	Charging Data Record
PDP	Packet Data Protocol
TCP	Transfer Control Protocol
UDP	User Datagram Protocol
URL	Uniform Resource Locator
WAP GW	Wireless Application Protocol Gateway
HTTP	Hyper Text Transfer Protocol
SGSN	Serving GPRS Support Node
LES	LAN Emulation Server
SMS	Short Message Service
MSISDN	Mobile Subscribe ISDN Number
QoS	Quality of Service
ToD	Time of Day
XML	Extensible Markup Language
SOAP	Simple Object Access Protocol
JS	Java Script
SS7	Signaling System #7
O&M	Operation & Maintenance
FTP	File Transfer Protocol
SNMP	Simple Network Management Protocol
CAI3	Common Air Interface 3G
CD	Common Directory
LDAP	Lighweight Directory Access Protocol
API	Application Programming Interface
CORBA	Common Object Request Broker Architecture
ISDN	Integrated Services Digital Network
UMTS	Universal Mobile Telecommunications System